# Model Order Reduction and Proper Orthogonal Decomposition

Ameir

**CNRS** 

March 19, 2024

#### Table of Contents

- Introduction to Model Order Reduction

  Motivation for MOR
  - Techniques in Model Order Reduction
  - Proper Orthogonal Decomposition (POD)
  - Applying POD to Fluid Dynamics
  - Application of POD
  - Step 1 Get Dataset
  - Step 2 Standardization
  - Step 3 Construct Correlation Matrix
  - Step 4 Eigenvalue Decomposition
  - Step 5 Obtain Reduced Dataset Row-Wise
  - Step 6 Column Correlation
  - Step 7 Eigenvalue Decomposition of Column Correlations
  - Step 8 Obtain Reduced Dataset Row-Wise and Column-Wise Correlated
  - Step 9 Compute Inverse
  - Step 10 Compare Reconstruction with Original (MSE)
- 5 Conclusio

#### Introduction to Model Order Reduction

- Model Order Reduction (MOR) is a technique used in numerical simulations to reduce the complexity of mathematical models.
- Essential for managing computational costs while preserving important features and dynamics.
- Key in engineering, physics, and applied sciences, especially with large-scale systems.

• Complexity and Cost of High-fidelity Models: High-fidelity models, essential for accurate simulations in engineering and physics, incorporate an immense amount of detail to replicate real-world phenomena accurately. Each additional detail, while improving model accuracy, also increases the computational resources required, including processing power and memory. For large-scale systems, such as those encountered in aerospace engineering, climate modeling, or structural analysis, the computational cost can become prohibitive, making simulations extremely slow and sometimes infeasible with available technology.

• Simplification through MOR: Model Order Reduction (MOR) techniques address these challenges by simplifying the original, high-dimensional models. MOR strategically reduces the number of equations or variables, focusing on preserving the model's core dynamics and behavior. This simplification process reduces computational demands significantly, enabling faster simulations without a corresponding loss in essential accuracy. The key lies in identifying and retaining the most influential features of the system while discarding the redundant or less significant aspects.

• Enabling Real-time and Interactive Applications: In many practical scenarios, such as control system design, interactive simulations for educational purposes, or real-time decision-making in engineering and environmental monitoring, the ability to perform quick simulations is crucial. Real-time simulation and analysis capabilities allow for immediate feedback and adjustments, which are vital in dynamic environments and systems where conditions change rapidly. MOR opens the door to these applications by making simulations sufficiently fast and responsive.

- **Design Optimization:** In the design and optimization of engineering systems, multiple iterations of simulations are often required to explore the effects of various parameters and configurations. MOR significantly accelerates this iterative process by reducing simulation times, thus enabling a more exhaustive exploration of the design space within practical time frames. This can lead to more innovative designs and optimizations that might not be feasible to explore with full-scale simulations due to time or resource constraints.
- The fundamental idea behind MOR is to create a simpler, reduced model that approximates the behavior of the original, high-fidelity model but with significantly fewer degrees of freedom. This is achieved by identifying the system's most critical modes or states that capture the majority of its dynamics, and then expressing the system's behavior in terms of these modes or states.

#### Table of Contents

- Introduction to Model Order Reduction

  Motivation for MOR
  - Techniques in Model Order Reduction
  - Proper Orthogonal Decomposition (POD)
  - Applying POD to Fluid Dynamics
  - Application of POD
  - Step 1 Get Dataset
  - Step 2 Standardization
  - Step 3 Construct Correlation Matrix
  - Step 4 Eigenvalue Decomposition
  - Step 5 Obtain Reduced Dataset Row-Wise
  - Step 6 Column Correlation
  - Step 7 Eigenvalue Decomposition of Column Correlations
  - Step 8 Obtain Reduced Dataset Row-Wise and Column-Wise Correlated
  - Step 9 Compute Inverse
  - Step 10 Compare Reconstruction with Original (MSE)
- 5 Conclusio

### Techniques in Model Order Reduction

- Proper Orthogonal Decomposition (POD): A powerful technique for identifying the most energetically significant modes of a system. By analyzing the covariance of the data, POD extracts orthogonal modes that represent the system's dominant patterns of variation. These modes can then be used to construct a reduced-order model that captures the essential dynamics of the original system with far fewer degrees of freedom. This approach is particularly useful in fluid dynamics and structural analysis, where it can significantly reduce the complexity of simulations.
- Krylov Subspace Methods: These methods approximate the solution of large linear systems by projecting them onto a smaller subspace, generated by the action of the system's matrix on an initial vector. Techniques such as the Arnoldi and Lanczos algorithms are used to construct a basis for this subspace, enabling efficient computations and analysis. Krylov subspace methods are widely applied in solving sparse systems of linear equations and eigenvalue problems.

### Techniques in Model Order Reduction

- Balanced Truncation: This technique focuses on reducing the system while maintaining a balance between its controllability and observability. It involves computing singular value decompositions of controllability and observability Gramians, identifying states that contribute least to the system's input-output behavior, and truncating them. Balanced truncation is particularly favored in control theory for its ability to preserve the essential input-output characteristics of the system.
- Moment Matching: Moment matching approximates the original system by ensuring that the reduced-order model matches the first few moments (i.e., derivatives at a specific point) of the system's transfer function. This approach is especially useful in circuit simulation and control systems, where the goal is to replicate the input-output response of the system accurately. By matching moments, the technique ensures that the reduced model behaves similarly to the original system under specific conditions.

#### Table of Contents

- Introduction to Model Order Reduction

  Motivation for MOR
  - Techniques in Model Order Reduction
    - Proper Orthogonal Decomposition (POD)
    - Applying POD to Fluid Dynamics
    - Application of POD
    - Step 1 Get Dataset
    - Step 2 Standardization
    - Step 3 Construct Correlation Matrix
    - Step 4 Eigenvalue Decomposition
    - Step 5 Obtain Reduced Dataset Row-Wise
    - Step 6 Column Correlation
    - Step 7 Eigenvalue Decomposition of Column Correlations
    - Step 8 Obtain Reduced Dataset Row-Wise and Column-Wise Correlated
    - Step 9 Compute Inverse
    - Step 10 Compare Reconstruction with Original (MSE)
- 5 Conclusio

# Understanding Proper Orthogonal Decomposition (POD)

- Principle of POD: Proper Orthogonal Decomposition (POD) is a statistical method used to reduce the dimensionality of a system by identifying patterns in the data. It decomposes a dataset into a set of orthogonal modes, each representing a direction of maximum variance. This process is akin to finding the principal axes in the data space along which the data variability is maximized.
- Projection onto Lower-dimensional Space: Once the significant modes (also known as eigenfunctions or singular vectors) are identified, the high-dimensional system can be approximated by projecting it onto the subspace spanned by these modes. This projection effectively captures the most critical aspects of the system's dynamics, allowing for a simplified model that retains the essence of the original system's behavior.

# Understanding Proper Orthogonal Decomposition (POD)

- Reduction of Variables: The process significantly reduces the number of variables needed to describe the system. For example, a fluid dynamics simulation with millions of data points (representing velocity and pressure across a spatial domain) can be approximated using a much smaller set of POD modes. This reduction facilitates faster simulations and analyses while using substantially less computational resources.
- Application and Benefits: POD is widely used in various fields such as fluid dynamics, image processing, and data analysis. Its ability to extract dominant features from complex datasets makes it an invaluable tool for simplifying models, enhancing computational efficiency, and identifying underlying structures in the data.

# Understanding Proper Orthogonal Decomposition (POD)

 Mathematical Foundation: Mathematically, POD involves solving an eigenvalue problem derived from the covariance matrix of the dataset or directly applying singular value decomposition (SVD) to the data matrix. The resulting eigenvalues and eigenvectors (or singular values and singular vectors) provide a hierarchy of modes ordered by their energy content or variance explained.

# Step 1: Organize the Dataset

- Data Preparation: Begin by structuring your CFD simulation data.
   Each data point should represent a snapshot in time, with variables X, Y, Z for spatial coordinates, and corresponding values for Velocity, Pressure, and Turbulent Viscosity.
- Normalization: Consider normalizing the data if the variables are on vastly different scales to ensure that no single variable dominates the POD due to its scale.

 POD Application: Apply POD to the dataset to extract significant modes. This involves computing the covariance matrix of the dataset and finding its eigenvalues and eigenvectors, or using Singular Value Decomposition (SVD) for a more direct approach.

#### Computing the Covariance Matrix:

- The first step in applying POD is to compute the covariance matrix of the dataset. This matrix captures the variance and covariance of the variables across the dataset, essentially quantifying how much the variables change together.
- 2 For a dataset with variables X, Y, Z, Velocity, Pressure, and Turbulent Viscosity, the covariance matrix provides a measure of the variability of these quantities with respect to each other, across all spatial points and time steps considered in the dataset.
- The covariance matrix,

$$\Sigma = \begin{bmatrix} \mathsf{Var}(X_1) & \mathsf{Cov}(X_1, X_2) & \cdots & \mathsf{Cov}(X_1, X_p) \\ \mathsf{Cov}(X_2, X_1) & \mathsf{Var}(X_2) & \cdots & \mathsf{Cov}(X_2, X_p) \\ \vdots & \vdots & \ddots & \vdots \\ \mathsf{Cov}(X_p, X_1) & \mathsf{Cov}(X_p, X_2) & \cdots & \mathsf{Var}(X_p) \end{bmatrix}$$

For each component,  $\text{Cov}(X_i, X_j) = \frac{1}{n-1} \sum_{k=1}^n (X_{ik} - \bar{X}_i)(X_{jk} - \bar{X}_j)$ 

#### Eigenvalue Decomposition:

- The next step involves eigenvalue decomposition of the covariance matrix. This process identifies the eigenvalues and their corresponding eigenvectors of the matrix, which represent the variance captured by each mode and the direction of these modes, respectively.
- The eigenvectors (modes) are orthogonal to each other, meaning they capture independent features or dynamics within the dataset. The eigenvalues indicate the amount of variance captured by each mode, providing a basis for selecting the most significant modes.

#### Using Singular Value Decomposition (SVD):

- ① Alternatively, SVD can be applied directly to the data matrix  $\mathbf{X}$  without explicitly computing the covariance matrix. SVD decomposes  $\mathbf{X}$  into three matrices,  $\mathbf{X} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T$ , where  $\mathbf{U}$  and  $\mathbf{V}$  are orthogonal matrices containing the left-singular vectors and right-singular vectors, and  $\mathbf{\Sigma}$  is a diagonal matrix with singular values.
- 2 The columns of U (the left-singular vectors) can be interpreted as the modes in POD, and the singular values in Σ represent the strength (or energy) of each mode. This method is numerically stable and efficient, especially for large datasets.

#### Mode Selection:

- After identifying the modes and their corresponding strengths, select a subset of modes that capture a significant portion of the system's energy. This is typically done by choosing modes associated with the largest singular values (or eigenvalues), which effectively capture the most important dynamics of the fluid flow.
- Mode Selection: Select a number of leading modes that capture a substantial portion of the system's energy or variance, which represent the dominant flow structures and dynamics.

## Step 3: Construct the Reduced Model

- Projection: Project the original dataset onto the subspace spanned by the identified significant modes. This step reduces the dimensionality from potentially millions of points to just a few modes.
- Model Formulation: The reduced model is formulated in terms of these modes, significantly simplifying the representation of the fluid dynamics.

# Step 4: Analyze and Simulate

- Simulation with Reduced Model: Use the reduced model to perform fluid dynamics analyses and simulations. This should be much faster and less resource-intensive than using the full-scale model.
- Interactive Exploration: The efficiency of the reduced model allows for interactive exploration of different scenarios and parameters, facilitating deeper insights into the fluid dynamics.

# Step 5: Validate the Reduced Model

- Comparison with Full Model: Validate the reduced model by comparing its predictions with those from the original full-scale model. This could involve comparing summary statistics, flow patterns, or specific dynamic behaviors.
- Adjustments: Based on this comparison, adjustments to the model or the selection of modes may be necessary to improve accuracy and representation.

#### Table of Contents

- Introduction to Model Order Reduction

  Motivation for MOR
  - Tachniques in Madel Order Reduction
  - Proper Orthogonal Decomposition (POD)
  - Applying POD to Fluid Dynamics
  - Application of POD
  - Step 1 Get Dataset
  - Step 2 Standardization
  - Step 3 Construct Correlation Matrix
  - Step 4 Eigenvalue Decomposition
  - Step 5 Obtain Reduced Dataset Row-Wise
  - Step 6 Column Correlation
  - Step 7 Eigenvalue Decomposition of Column Correlations
  - Step 8 Obtain Reduced Dataset Row-Wise and Column-Wise Correlated
  - Step 9 Compute Inverse
  - Step 10 Compare Reconstruction with Original (MSE)
- Conclusion

#### Get Dataset

- Get dataset -
- Define dataset  $X : X = [x_{ij}] = \{\vec{X}_i(\theta_j), \vec{V}_i(\theta_j) | \forall i \in [1, N], \forall j \in [1, n \times M]; i, j \in \mathbb{N}\}$
- In our example,  $X \in \mathbb{R}^{N \times 2M} = \mathbb{R}^{20 \times 12}$
- We take N=20 (20 observations), n=2 (2 angles), and M=6 (6 variables)

# Dataset $X \in \mathbb{R}^{20 \times 12}$

Points:0 ( $\theta = 0$ )	Points: $1 (\theta = 0)$	Points:2 ( $\theta = 0$ )	Velocity:0 ( $\theta = 0$ )	Velocity:1 ( $\theta = 0$ )	Velocity:2 ( $\theta = 0$ )	Points:0 ( $\theta = 135$ )	Points:1 ( $\theta = 135$ )	Points:2 ( $\theta = 135$ )	Velocity:0 ( $\theta = 135$ )	Velocity:1 ( $\theta = 135$ )	Velocity:2 ( $\theta = 135$ )
-49.809	-233.39	1.1732	-0.11436	-0.16909	-0.012407	-49.809	-233.39	1.1732	-0.28844	-0.50923	-0.063486
162.71	-33.957	20.784	0.29694	-0.11781	-0.053469	162.71	-33.957	20.784	-0.14046	-0.0409	0.23128
-577.8	270.26	1.7004	-0.28844	-0.88723	0.0021901	-577.8	270.26	1.7004	-0.47265	0.69779	-0.032282
-409.59	635.01	27.039	-0.45833	-2.0245	0.032774	-409.59	635.01	27.039	-0.70382	0.099895	0.11435
-45.52	-226.9	15.933	0.016559	-0.020085	-0.038222	-45.52	-226.9	15.933	-0.43829	-0.32702	-0.41848
187.5	-50.85	49.567	0.26444	-0.65064	-0.23518	187.5	-50.85	49.567	-0.62553	0.79691	0.11876
400.15	-712.43	23.415	-0.28017	-2.3111	-0.042736	400.15	-712.43	23.415	-1.6547	1.63	-0.00031295
251.73	-610.99	219.74	-0.031567	-3.6066	-0.28337	251.73	-610.99	219.74	-3.3551	3.285	0.044889
44.669	-98.604	56.746	0.022383	0.17762	-0.0021013	-44.669	-98.604	56.746	-0.47375	0.42469	0.023411
-142.85	-160.2	7.2621	-0.034768	-0.082606	-0.017066	-142.85	-160.2	7.2621	-0.49406	0.12969	0.020166
-56.055	-108.65	49.824	0.030558	0.11737	0.0087919	-56.055	-108.65	49.824	0.016165	0.1787	0.12693
198.23	-726.9	234.14	0.020346	-3.7919	-0.26086	198.23	-726.9	234.14	-3.4149	3.3526	0.034918
-407.78	638.38	29.006	-0.44665	-2.0776	0.033642	-407.78	638.38	29.006	-0.73376	0.10373	0.12352
826.64	-268.21	49.689	0.044228	-3.489	0.0025637	826.64	-268.21	49.689	-2.2013	2.2569	0.022693
386.76	928.29	27.776	0.17297	-2.6085	0.069452	386.76	928.29	27.776	-1.9341	2.0256	0.0031161
234.02	68.108	216.36	-0.32819	-1.4283	0.14909	234.02	68.108	216.36	-3.1677	3.2374	0.35045
14.569	15.468	153.18	-0.22886	-2.1073	0.7425	-14.569	15.468	153.18	0.68777	0.26538	0.12476
6.6665	12.702	136.78	-0.054038	-0.52924	-0.26808	-6.6665	12.702	136.78	0.7487	0.04379	0.23887
474.07	435.17	31.053	-0.96927	-2.0113	0.061435	-474.07	435.17	31.053	-0.15149	0.072373	-0.0096399
173.3	42 965	55 431	.0.20574	.1 1827	0.5018	173.3	42 965	55 431	.0.81865	1.21	0.17383

#### Standardize Dataset

• Define  $f: \mathbb{R}^{20 \times 12} \to \mathbb{R}^{20 \times 12}$  such that  $f(X) = Z = [z_{ij}]$ , where  $Z \in \mathbb{R}^{N \times M}$  is the standardized matrix, and each element  $z_{ij}$  of Z is given by:

$$z_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j}$$

- With  $\mu_j = \frac{1}{N} \sum_{i=1}^{N} x_{ij}$  is the mean of the *j*-th column of X,
- And  $\sigma_j = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_{ij} \mu_j)^2}$  is the standard deviation of the *j*-th column of X,
- $x_{ij}$  represents the original value in the *i*-th row and *j*-th column of X,
- $z_{ij}$  represents the standardized value corresponding to  $x_{ij}$  in Z.
- This function f thus transforms each feature in X to have a mean of 0 and a standard deviation of 1 in Z, applied column-wise across the entire matrix.



# Standardized Dataset $Z \in \mathbb{R}^{20 \times 12}$

Points:0 (θ = 0)	Points:1 (θ = 0)	Points:2 (θ = 0)	Velocity:0 (θ = 0)	Velocity:1 (θ = 0)	Velocity: $2 (\theta = 0)$	Points:0 (θ = 135)	Points:1 (θ = 135)	Points:2 (θ = 135)	Velocity:0 (θ = 135)	Velocity:1 (θ = 135)	Velocity.2 ( $\theta = 135$ )
-0.2422775230642614	-0.5299456452302135	-0.9263562703310546	0.05045384463351536	1.002311924281608	-0.13555268487255895	-0.2422775230642614	-0.5299456452302135	-0.9263562703310546	0.5732341608441163	-1.1942511144808337	-0.8349562911482649
0.4062579500808071	-0.05844463045642818	-0.6636690924848442	1.5079436024824224	1.0427534364237152	-0.30979510525014087	0.4062579500808071	-0.05844463045642818	-0.6636690924848442	0.695752493701489	-0.8100862791999293	1.135991818641859
-1.8535258961571983	0.6607875137295223	-0.9192944125710181	-0.5664190390191414	0.4359572532649747	-0.07361137576697227	-1.8535258961571983	0.6607875137295223	-0.9192944125710181	0.42071961769195637	-0.20414875187165366	-0.6263201892983533
-1.3402063720782595	1.5231322367128712	-0.5798832042279629	-1.1684441665682561	-0.4609405336666635	0.05616829304082875	-1.3402063720782595	1.5231322367128712	-0.5798832042279629	0.22932508460208717	-0.6945940237745349	0.35413729071883119
-0.22918895853911686	-0.5146019378904961	-0.7286483640970153	0.5143806659508533	1.1198233813162162	-0.24509601159914984	-0.22918895853911686	-0.5146019378904961	-0.7286483640970153	0.44916758307227345	-1.0447866884638508	-3.2086399958049596
0.48190857133945336	-0.09838318948413953	-0.2781200562421884	1.392776043205737	0.6225418351234219	-1.0808672510109605	0.48190857133945336	-0.09838318948413953	-0.2781200562421884	0.2941443863955978	-0.12284194156045591	0.3836248342760069
1.1308438117852433	-1.662495655094349	-0.6284267788151632	-0.5371133247049901	-0.6869650535560543	-0.26425071145752027	1.1308438117852433	-1.662495655094349	-0.6284267788151632	-0.5579450353291099	0.5605306432070555	-0.412558566995286
0.677916615044288	-1.4226704359046825	2.0013517203106885	0.3438405441754181	-1.7086495110463045	-1.2853566143638335	0.677916615044288	-1.4226704359046825	2.0013517203106885	-1.9657715875028936	1.9181050082699318	-0.1103149231190327
-0.22659199691382	-0.21128355933018983	-0.18195716386870311	0.5350186925730561	1.2757416566293076	-0.09182149516710755	-0.22659199691382	-0.21128355933018983	-0.18195716386870311	0.4198088853893269	-0.42816902927266365	-0.2539279527156357
-0.5262069079987394	-0.3569092908089625		0.3324974254910652	1.0705167553287611	-0.1553226774822505	-0.5262069079987394	-0.3569092908089625	-0.8447952959094199	0.402993455328959	-0.670153583770276	-0.2756257030201788
-0.26133818382995344	-0.23503438890411912	-0.27467753453443167	0.5639877632524016	1.228226034426461	-0.04559730896923422	-0.26133818382995344	-0.23503438890411912	-0.27467753453443167	0.8254283545190724	-0.6299513362586557	0.4382536848116362
0.5146528700931958	-1.6967057391013924	2.194240096153868	0.5278003443193771	-1.8547846943523203	-1.1898377189690799	0.5146528700931958	-1.6967057391013924	2.194240096153868	-2.0152823072276598	1.9735563841480286	-0.1769861922366287
-1.3346828696789796	1.5310996163330326	-0.5535351878888286	-1.1270547175731789	-0.5028173724877611	0.05985156275218575	-1.3346828696789796	1.5310996163330326	-0.5535351878888286	0.20453660720142672	-0.6914482245660659	0.4154526590678791
2.4323457666299406	-0.6122673539465403	-0.2764858630579614	0.6124290104923592	-1.6159053256346074	-0.07202604216770618	2.4323457666299406	-0.6122673539465403	-0.2764858630579614	-1.010496195890266	1.074768328680117	-0.2587288725056701
1.0899821005657082	2.2165070423511346	-0.5700110699921003	1.0686413771468306	-0.9215068965954986	0.2118076553370868	1.0899821006657082	2.2165070423511346	-0.5700110699921003		0.8850360322214808	-0.3896301630871832
0.6238717379773003	0.18285821848475212	1.9560765320919424	-0.7072778230563261	0.009247249857603246	0.5497434079558065	0.6238717379773003	0.18285821848475212	1.9560765320919424	-1.810615920673108	1.8790593649001475	1.9328241464669438
-0.13473706751142983	0.05840633029326893	1.1097787830799912	-0.3552903257377462	-0.5262400111504092	3.0678183842419315		0.05840633029326893	1.1097787830799912	1.3814759618894994	-0.5588488915676285	0.4237439411565179
-0.11062133395982224	0.05186693206497024	0.8901003550363696	0.2642119218836048	0.7182828564582859	-1.2204750545865886	-0.11062133396982224	0.05186693206497024	0.8901003550363696	1.4319222520706032	-0.7406162048934791	1.1867424888457054
-1.5369773304791274	1.0505689894449002	-0.5261155694616766	-2.979019943070373	-0.45053047203881974	0.17778833128645424	-1.5369773304791274	1.0505689894449002	-0.5261155694616766	0.6866203325214861	-0.717169952279224	-0.474923384315149

#### Construct Correlation Matrix

- Our goal is to reduce the number of observations.
- Therefore, we construct the correlation matrix as  $C_R = Z \cdot Z^T \in \mathbb{R}^{20 \times 20}$ ;  $Z = [z_{ij}], Z^T = [z'_{ik}]$
- The subscript R in  $C_R$  makes it clear that this correlation is done row wise.
- The construction of the correlation matrix can be defined by bilinear function  $g_R: \mathbb{R}^{20 \times 12} \times \mathbb{R}^{12 \times 20} \to \mathbb{R}^{20 \times 20}$  where each element  $(c_R)_{ik}$  of  $C_R$  is given by

$$(c_R)_{ik}=z_{ij}*z'_{jk}$$

# Row Correlation Matrix, $\mathrm{C}_R \in \mathbb{R}^{20 \times 20}$

5.X739061986427389			7.173 MINITERSTY	1 200 KIND TO DESIGN	LOCATION CONTROL		2 96299 (1299-67639)	T STATES AND PARTY.		7 ATTENCTORESCENCE	E-CINACIDAL CONONCRET	3.3616.60636613.304					1 ADMIGRATION IN	© E751 BHOSELEW/11
247WTH425/D039	1001722213000	-14DTHHOUSEHUS		4.196609366200743	£30004444707WEE	125/2003/24E/42T					-1.50mm256.07.290071		620000C0000010		1.738994/96968144	1140411401441064	-3-171181009634840	THE SHORT STATES
		\$12003/500360X1	£133900023425003	193090120312068				1.80/201301460/21901	1 PROBERTO PROPERTY		E-110100211200002	DOSESSED HITESON	1.368000360136	TERRORISTERIOR		E MREDOMOTIONOS	30 24 E11 E11 E0 COST	T BERSEITHBUNG

# Eigenvalue Decomposition of $C_R$

- ullet  $\mathrm{C}_R$  and indeed, any correlation matrix is a real and symmetric matrix.
- Therefore, the decomposition must result in real valued eigenvalues and eigenvectors.
- Solve the characteristic equation to get the eigenvalues  $det(C_R \lambda I) = 0$
- Use the matrix of eigenvalues  $\Lambda \in \mathbb{R}^{20 \times 20}$  to determine the matrix of eigenvectors  $V \in \mathbb{R}^{20 \times 20}$
- Mathematica or Matlab is effective for this.

# Eigenvalue Decomposition of $C_R$

- Sort  $\Lambda = \text{diag}(\lambda_1, \lambda_2, ..., \lambda_{20}), \in \mathbb{R}^{20 \times 20}$  such that  $\lambda_i > \lambda_j, \forall i < j$
- ullet Sort the eigenvectors  $V \in \mathbb{R}^{20 imes 20}$  according to their position in  $\Lambda$

# Sorted Eigenvalues and Sorted Eigenvectors $\Lambda, V \in \mathbb{R}^{20 \times 20}$



Table: Sorted Eigenvalues  $\Lambda \in \mathbb{R}^{20 \times 20}$ 



Table: Sorted Eigenvalues  $V \in \mathbb{R}^{20 \times 20}$ 

# Deleting Eigenvalues

- Construct  $\Lambda' = diag(\lambda_1, \lambda_2, ..., \lambda_k)$
- k is chosen such that it captures most of the variance present in Z as well as satisfying 1 < k < min(N, nM) (by a fundamental theorem in Linear Algebra).
- Here, we take k = 12 as the obvious choice.
- $\bullet$  Similarly, construct  $V^{\prime}$  by deleting the eigenvectors corresponding to the deleted eigenvalues.
- $\Lambda' \in \mathbb{R}^{k \times k}$  and  $V' \in \mathbb{R}^{N \times k}$  where N=20 and k=12.

# Reduced Eigenvalues and Reduced Eigenvectors $\Lambda' \in \mathbb{R}^{12 \times 12}$ . $V' \in \mathbb{R}^{20 \times 12}$

0	1	2	3	4	5	6	7	8	9	10	11
104.33544793460241		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	45.25316341800748		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	28.644758524673954		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	24.726725619237783	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	18.543985544275134		0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	9.102650911223483	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	5.580315312915288		0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.18519644217231	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.6277562928916468		0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.1968162656875224e-14	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.0144458408994026e-16	
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-9.692352108704102e-17

Table: Reduced Eigenvalues  $\Lambda \in \mathbb{R}^{12 \times 12}$ 

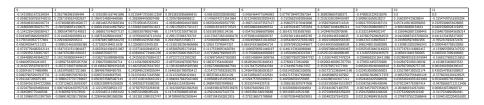


Table: Reduced Eigenvectors  $\boldsymbol{\Lambda} \in \mathbb{R}^{20 \times 12}$ 

# Obtaining Reduced Dataset Row-Wise

- By projecting  $(V')^T$  onto the original dataset Z, we obtain the reduced dataset row-wise.
- Define bilinear function  $h: \mathbb{R}^{12 \times 20} \times \mathbb{R}^{20 \times 12} \to \mathbb{R}^{12 \times 12}$  such that  $h\left((V')^T, \mathbf{Z}\right) = (V')^T \mathbf{Z} = \mathbf{W} \in \mathbb{R}^{12 \times 12}$
- ullet W is the new row reduced dataset. In comparison to Z, it has 8 less rows but the same number of columns with the same column-names.

# Reduced Dataset Row-Wise $\mathbf{W} \in \mathbb{R}^{12 \times 12}$



#### Column Correlation

- Having obtained  $W \in \mathbb{R}^{12 \times 12}$  it seems that we are ready to train our neural network.
- Keeping it mind that our goal is to predict velocity at new angles not seen in the dataset, it is probably best to correlate the columns as well.
- However, this time, we do not have to reduce the number of columns and we can keep the maximum number of columns present such that the full variance is captured by the eigenvalues, small as they may be.
- We repeat the process described from Step 3 to Step 5 with a few key changes.

#### Column Correlation

- We construct the correlation matrix as  $C_C = W^T \cdot W \in \mathbb{R}^{12 \times 12}; W = [w'_{jk}], W^T = [w_{ij}]$
- The subscript C in C<sub>C</sub> makes it clear that this correlation is done column wise.
- The construction of the correlation matrix can be defined by bilinear function  $g_C: \mathbb{R}^{12 \times 12} \times \mathbb{R}^{12 \times 12} \to \mathbb{R}^{12 \times 12}$  where each element  $(c_C)_{ik}$  of  $C_C$  is given by

$$(c_C)_{ik} = w_{ij} * w'_{jk}$$

# Column Correlation Matrix, $C_C \in \mathbb{R}^{12 \times 12}$

0	1	2	3	4	5	6	7	8	9	10	11
19.99999999999993	-9.088581217770104	5.940235106990982	11.418161016264317	-7.514357771780367	-1.49996778948842	19.99999999999993	-9.088581217770104	5.940235106990982	-10.489954061881173	12.042233349929832	2.1946170392122206
-9.088581217770104	19.99999999999982	-7.730034573127007	-6.558619238143711	1.8728995443158716	5.992601051051302	-9.088581217770104	19.99999999999982	-7.730034573127007	6.484088679730555	-6.703770197901935	3.271619071725043
5.940235106990982	-7.730034573127007	20.0	0.9638920777909981	-9.754405795273739	-1.2535602826043506	5.940235106990982	-7.730034573127007	20.0	-11.076742826768315	13.975395341146964	7.900862067615191
11.418161016264317	-6.558619238143711	0.9638920777909981	20.0000000000000004	4.140608143468153	-5.959438424719385	11.418161016264317	-6.558619238143711	0.9638920777909981	-2.0100967176547075	2.6643866024357803	-0.432220677551209
-7.514357771780367	1.8728995443158716	-9.754405795273739	4.140608143468153	20.0	0.5213289040614937	-7.514357771780367	1.8728995443158716	-9.754405795273739	13.784719493126806	-14.475521586391638	-1.1413928249713878
-1.49995778948842	5.992601051051302	-1.2535602826043506		0.5213289040614937	19.9999999999999	-1.49996778948842	5.992601051051302	-1.2535602826043506	6.003662279305754	-3.405280050860079	3.0358500965194507
19.99999999999993	-9.088581217770104	5.940235106990982	11.418161016264317	-7.514357771780367	-1.49996778948842	19.99999999999993	-9.088581217770104	5.940235106990982	-10.489954061881173	12.042233349929832	2.1946170392122206
-9.088581217770104	19.99999999999982	-7.730034573127007	-6.558619238143711	1.8728995443158716	5.992601051051302	-9.088581217770104	19.99999999999982	-7.730034573127007	6.484088679730555	-6.703770197901935	3.271619071725043
5.940235106990982	-7.730034573127007	20.0	0.9638920777909981	-9.754405795273739	-1.2535602826043506	5.940235106990982	-7.730034573127007	20.0	-11.076742826768315	13.975395341146964	7.900862067615191
-10.489954061881173	6.484088679730555	-11.076742826768315	-2.0100967176547075	13.784719493126806	6.003662279305754	-10.489954061881173	6.484088679730555	-11.076742826768315	20.0	-18.653399619727043	-1.0372414462435116
12.042233349929832	-6.703770197901935	13.975395341146964	2.6643866024357803	-14.475521586391638	-3.406280060860079	12.042233349929832	-6.703770197901935	13.975395341146964	-18.653399619727043	20.0000000000000001	4.598221421175162
2.1946170392122206	3.271619071725043	7.900862067615191	-0.432220677551209	-1.1413928249713878	3.0358600965194507	2.1946170392122206	3.271619071725043	7.900852057615191	-1.0372414462435116	4.598221421175162	19.99999999999999

# Eigenvalue Decomposition of $\mathrm{C}_{\mathcal{C}}$

- Again, the decomposition must result in real valued eigenvalues and eigenvectors.
- Solve the characteristic equation to get the eigenvalues  $det(C_C \lambda I) = 0$
- Use the matrix of eigenvalues  $\Lambda_C \in \mathbb{R}^{12 \times 12}$  to determine the matrix of eigenvectors  $V_C \in \mathbb{R}^{12 \times 12}$

# Eigenvalue Decomposition of $\mathrm{C}_{\mathcal{C}}$

- Sort  $\Lambda = \text{diag}(\lambda_1, \lambda_2, ..., \lambda_{12}), \in \mathbb{R}^{12 \times 12}$  such that  $\lambda_i > \lambda_j, \forall i < j$
- Sort the eigenvectors  $V_C \in \mathbb{R}^{12 \times 12}$  according to their position in  $\Lambda$
- Construct  $\Lambda_C' = diag(\lambda_1, \lambda_2, ..., \lambda_{k'})$
- Since we want all our columns to stay, we want to take k'=12.
- ullet We simply use  $\Lambda_C \in \mathbb{R}^{12 imes 12}$  and  $V_C \in \mathbb{R}^{12 imes 12}$

# Sorted Eigenvalues and Sorted Eigenvectors $\Lambda_C, V_C \in \mathbb{R}^{12 \times 12}$

0	1	2	3	4	5	6	7	8	9	10	11
104.3354479346024		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	45.25316341800758	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	28.64475852467392			0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	24.726725619237776	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	18.543985544275124		0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	9.102650911223483	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	5.580315312915283	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.1851964421723067		0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.6277562928916579	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0658856570233323e-14	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.302915155099587e-15	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-8.168869046324322e-16

Table: Sorted Eigenvalues  $\Lambda \in \mathbb{R}^{12 \times 12}$ 

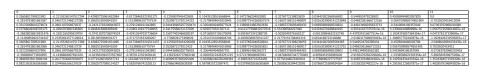


Table: Sorted Eigenvalues  $V \in \mathbb{R}^{12 \times 12}$ 

#### Obtaining Reduced Dataset Row-Wise and Column-Wise

- Now, we project the reduced dataset W onto the eigenvectors ( $V_C$ ) to obtain the reduced dataset column-wise.
- Define bilinear function  $h_C: \mathbb{R}^{12 \times 12} \times \mathbb{R}^{12 \times 12} \to \mathbb{R}^{12 \times 12}$  such that  $h_C(W, V_C) = WV_C = W_C \in \mathbb{R}^{12 \times 12}$
- ullet  $W_{\mathcal{C}}$  is the new column reduced dataset. In comparison to W, it has the same number of rows and columns but different values.
- ullet Now, we can use  $W_{\mathcal{C}}$  for our training dataset.

# Reduced Datasets $W, W_{C} \in \mathbb{R}^{12 \times 12}$



Table: Reduced Dataset W



Table: Reduced Dataset W<sub>C</sub>

- Now, importantly, we need to know how we can reconstruct the original dataset.
- We first summarize the functions we took to get to the final reduced dataset  $W_C \in \mathbb{R}^{12 \times 12}$ .
- $\mathcal{F}(X) = h_C \circ g_{\lambda_C} \circ g_C \circ h_R \circ g_{\lambda_R} \circ g_R \circ f(X) = W_C$
- I did not explicitly define  $g_{\lambda_C}: \mathbb{R}^{20 \times 20} \to \mathbb{R}^{20 \times 12}, g_{\lambda_R}: \mathbb{R}^{12 \times 12} \to \mathbb{R}^{12 \times 12}$  but it is clear that they are the reduction process of the eigenvectors.
- This step is nuanced as there is a potential loss of information in this step.

- If one did not (or could not) retain all the variance, one loses information in this step that would be impossible to recover.
   Therefore, any reconstruction is necessarily an approximation.
- This loss of information might seem "non-linear" when used colloquially but mathematically, the function is still a linear one.
- $m{\mathcal{F}}$  is linear as all the individual components are linear and therefore  $\mathcal{F}^{-1}$  necessarily exists.
- Whether or not  $\mathcal{F}^{-1}$  gives a good reconstruction/approximation of the original dataset has to do with  $g_{\lambda_C}, g_{\lambda_R}$  and how much variance one was able to keep.

- It is somewhat trivial to see that  $(h_C \circ g_{\lambda_C} \circ g_C)^{-1} = \mathcal{T}_C : \mathbb{R}^{12 \times 5} \times \mathbb{R}^{12 \times 5} \to \mathbb{R}^{12 \times 5}$
- $\mathcal{T}_C(W_C, V_C') = W_C(V_C')^T = W_R$  where the subscript  $\mathcal{R}$  indicates that it is reconstructed.
- And similarly, we see that  $(h_R \circ g_{\lambda_R} \circ g_R)^{-1} = \mathcal{T}_R : \mathbb{R}^{12 \times 5} \times \mathbb{R}^{12 \times 5} \to \mathbb{R}^{12 \times 5}$
- $\mathcal{T}_R(V',W) = (V')W = Z_{\mathcal{R}}$  where the subscript  $\mathcal{R}$  indicates that it is reconstructed.

# Reconstruction of Dataset $Z_{\mathcal{R}} \in \mathbb{R}^{20 \times 12}$

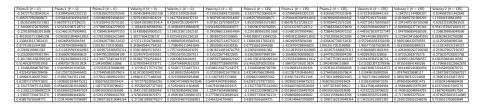


Table: Original Dataset Z

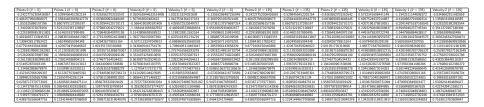


Table: Reconstructed Dataset  $Z_{\mathcal{R}}$ 

- In evaluating how good our approximation was, the best way is to compare  $Z_{\mathcal{R}}$  with Z
- We can do this via the Mean Square Errors method, defined as  $MSE(Z, Z_{\mathcal{R}}) = \frac{1}{mn} \sum_{i=1}^{m} \sum_{i=1}^{n} (z_{ij} z_{\mathcal{R}ij})^2$
- In our case, we obtain MSE  $\sim$  7 $E-31 \approx 0$

#### Table of Contents

- Introduction to Model Order R
  - Techniques in Model Order Reduction
  - Proper Orthogonal Decomposition (POD)
  - Applying POD to Fluid Dynamics
  - Application of POD
  - Step 1 Get Dataset
  - Step 2 Standardization
  - Step 3 Construct Correlation Matrix
  - Step 4 Eigenvalue Decomposition
  - Step 5 Obtain Reduced Dataset Row-Wise
  - Step 6 Column Correlation
  - Step 7 Eigenvalue Decomposition of Column Correlations
  - Step 8 Obtain Reduced Dataset Row-Wise and Column-Wise Correlated
  - Step 9 Compute Inverse
  - Step 10 Compare Reconstruction with Original (MSE)
- 5 Conclusion

#### Conclusion and Future Outlook

- Critical Role of MOR: Model Order Reduction, especially techniques like Proper Orthogonal Decomposition (POD), plays a pivotal role in tackling the computational challenges posed by high-fidelity models in science and engineering. By distilling complex systems to their most influential dynamics, MOR enables simulations that were once impractical due to computational constraints, making it an indispensable tool in the researcher's toolkit.
- Advancements in Computational Efficiency: The application of MOR methodologies, such as POD, has led to significant advancements in computational efficiency. This not only accelerates the simulation process but also opens up new possibilities for real-time analysis and control, interactive simulation environments, and the ability to handle larger and more complex models than ever before.

#### Conclusion and Future Outlook

- Wide-ranging Applications: The utility of MOR extends across numerous fields, including but not limited to fluid dynamics, structural analysis, electrical engineering, and control systems. Its versatility and effectiveness in simplifying complex problems while retaining essential characteristics make it a cornerstone technique in both theoretical research and practical applications.
- Future Directions: The ongoing development of MOR techniques promises to further broaden their applicability and enhance their effectiveness. Innovations in algorithmic efficiency, the integration of machine learning methods for automated mode selection, and the exploration of non-linear MOR methods are just a few areas where significant progress can be expected. As computational power continues to grow, the synergy between MOR techniques and high-performance computing will undoubtedly lead to new breakthroughs in simulation and modeling capabilities.

#### Conclusion and Future Outlook

 Final Thoughts: In conclusion, the strategic simplification offered by MOR, particularly through POD, represents a fundamental advance in our ability to understand and predict complex systems. As we continue to push the boundaries of what is computationally feasible, MOR will remain at the forefront of this endeavor, enabling us to tackle the grand challenges of science and engineering with ever-greater sophistication and insight.

## Questions?

Questions?